数理解析研究所講究録
第 833 巻 1993 年 214-224

正データからのEFSの帰納推論可能性について

—有限の弾力性をもつ言語族の Closure Properties とその特徴付け—

大阪府立大学総合科学部　森山隆史 (Takashi Moriyama)
大阪府立大学総合科学部　佐藤優子 (Masako Sato)

1 Abstract

In this paper, we consider a subfamily of language classes to be recursively inferable from positive data, which is closed under various operations such as sum, intersection, concatenation and so on. It is shown that the family of language classes $\boldsymbol{L}$ with finite elasticity is closed under such operations. It implies that any language class obtained by finitely applying operations to given ones contained in $\boldsymbol{L}$ is inferable from positive data.

The above discussions on closure operations are developed for an elementary formal system (EFS for short) as a unifying framework for inductive inference of various language classes. We introduce max length bounded EFSs by which any contex sensitive language can be represented, and for which operations corresponding to language operations can naturally be defined. It is shown that any EFS language class obtained by applying such operations to given max length bouded EFSs defining language classes with finite elasticity is inferable from positive data.

Furthermore, we obtain two theorems characterizing a max length bouded EFS language class to have finite elasticity, and to be inferable from positive data, respectively.

2 言語族の帰納推論

2.1 準備

$L_1, L_2, \cdots$ をアルファベット Σ 上の言語とする. 言語族 $\mathcal{L} = L_1, L_2, \cdots$ が帰納的言語の添字付き族であるとは, 任意の $i \in N$ と 任意の語 $w \in \Sigma^*$ に対して, $w \in L_i$ かどうかを決定する計算可能関数 f が存在することをいう. 以降, 言語族は帰納的言語の添字付き族とする. $\mathcal{L}$ の言語 L に対して, L のすべての語が少なくとも一度は現れるような無限列 $w_1, w_2, \cdots$ を言語 L の正提示という. 推論機械とは, ときどき入力を要求し, ときどき推測を生成するような実行的手続きである. 推論機械 M が 正データから L を推論するとは, L の任意の正提示を M に与えたとき, M の生成する推測の列が $L = L_j$ となる j に収束することをいう. 言語族 $\mathcal{L}$ が正データから帰納推論可能であるとは, 任意の $L \in \mathcal{L}$ を正データから推論する推論機械 M が存在することである.

定義 2.1 $\mathcal{L}$ を言語族, $L \in \mathcal{L}, T \subseteq \Sigma^*$ とする.

$T : \mathcal{L}$ で L の有限証拠集合 (以下, ftt)

$\iff$ (1) $T : L$ の有限部分集合, (2) $\nexists L' \in \mathcal{L}$ s.t. $T \subseteq L' \subsetneq L$.

$\mathcal{L}$: ftt をもつ $\iff \forall L \in \mathcal{L}, \exists T \subseteq \Sigma^*$ s.t. $T : \mathcal{L}$ で L の ftt.

Angluin [1] は, 言語族が正データから帰納推論可能であるための必要十分条件は, その任意の言語の ftt の要素を枚挙する実行的な手続きが存在することであることを示した. さらに, 次にのべる有限の厚さという条件が帰納推論可能であるための十分条件であることを示し, それを用いてパターン言語と呼ばれる言語族が正データから帰納推論可能であることを示した [1] .

定義 2.2 $\mathcal{L}$ を言語族とする.

$\mathcal{L}$: 有限の厚さ $\iff \forall w \in \Sigma^*, \#\{L \in \mathcal{L} \mid w \in L\} < \infty$

集合 $S \subseteq \Sigma^*$ が集合対 $I = (T, F)$ $(T, F \subseteq \Sigma^*)$ と無矛盾であるとは, $T \subseteq S$ かつ $F \subseteq S^C$ となることである.

定義 2.3 $\mathcal{L}$ を言語族, $L \in \mathcal{L}, I = (T, F)$ を有限な集合対とする.

$I : \mathcal{L}$ で L の有限証拠集合対 (以下, pftt)

$\iff$ (1) $L : I$ に無矛盾, (2) $\nexists L' \in \mathcal{L}$ s.t. $L \not\subseteq L'$ かつ $L' : I$ に無矛盾

pftt をもつ言語族は正データから帰納推論可能であることが示されている [4].

定義 2.4 $\mathcal{L}$ を言語族, $L \in \mathcal{L}, S \subseteq \Sigma^*$ とする.

$L : S$ の極小言語 $\iff S \subseteq L$ かつ $\exists L' \in \mathcal{L}$ s.t. $S \subseteq L' \subsetneq L$

空でない有限集合 $S \subseteq \Sigma^*$ に対して, $MIN(S, \mathcal{L}) = \{L \in \mathcal{L} \mid L : S$の極小言語 $\}$ とする.

定義 2.5 $\mathcal{L}$ を言語族とする.

$\mathcal{L}$: M - 有限の厚さ $\iff \forall S (\subseteq \Sigma^*)$: 有限集合, $\#MIN(S, \mathcal{L}) < \infty$

定理 2.1 [3] 言語族 $\mathcal{L}$ を有限の厚さとする. 次の (1), (2), (3) は同値である:

(1) $\mathcal{L}$: 正データから帰納推論可能, (2) $\mathcal{L}$: ftt をもつ, (3) $\mathcal{L}$: pftt をもつ

2.2 有限な弾力性をもつ族の Closure Properties

Wright [7] は有限の厚さを一般化した有限の弾力性という概念を導入し，この性質が正データから帰納推論可能であるための十分条件であることを示した．この節では言語族の演算 (和, 共通部分, 連接等) を導入し，有限の弾力性という言語族の性質がこれらの演算に関して保持されることを示す．

定義 2.6 $\mathcal{L}$ を言語族とする．

$\mathcal{L}$: 有限の弾力性をもつ $\iff \not\exists w_0, w_1, \cdots (\in \Sigma^*),\ \not\exists L_1, L_2, \cdots (\in \mathcal{L})$ s.t.

$$\forall k \in N,\ \{w_0,\ w_1,\ \cdots,\ w_{k-1}\} \subseteq L_k,\ w_k \notin L_k$$

定理 2.2 [7] 有限の弾力性をもつ言語族は正データから帰納推論可能である．

有限の弾力性をもつ言語族のすべての集まりを $\boldsymbol{L}$ とおく．

言語族 $\mathcal{L}_1, \mathcal{L}_2$ の 和演算 $\tilde{\cup}$, 共通部分演算 $\tilde{\cap}$, 連接演算 $\tilde{\cdot}$ を次のように定義する:

$$\begin{aligned}
\mathcal{L}_1 \tilde{\cup} \mathcal{L}_2 &= \{L_1 \cup L_2 \mid L_1 \in \mathcal{L}_1, L_2 \in \mathcal{L}_2\} \\
\mathcal{L}_1 \tilde{\cap} \mathcal{L}_2 &= \{L_1 \cap L_2 \mid L_1 \in \mathcal{L}_1, L_2 \in \mathcal{L}_2\} \\
\mathcal{L}_1 \tilde{\cdot} \mathcal{L}_2 &= \{L_1 \cdot L_2 \mid L_1 \in \mathcal{L}_1, L_2 \in \mathcal{L}_2\}
\end{aligned}$$

定理 2.3 [7] 言語族 $\mathcal{L}_1, \mathcal{L}_2$ が有限の弾力性をもつならば，$\mathcal{L}_1 \tilde{\cup} \mathcal{L}_2$ も有限の弾力性をもつ．

定理 2.4 言語族 $\mathcal{L}_1, \mathcal{L}_2$ が有限の弾力性をもつならば，$\mathcal{L}_1 \tilde{\cap} \mathcal{L}_2$, $\mathcal{L}_1 \tilde{\cdot} \mathcal{L}_2$, $\mathcal{L}_1 \cup \mathcal{L}_2$ も有限の弾力性をもつ．

proof. ここでは $\mathcal{L}_1 \tilde{\cdot} \mathcal{L}_2$ についてのみ証明する．任意の $k \in N$ に対して，

$$\{w_0, w_1, \cdots w_{k-1}\} \subseteq L_k^1 \cdot L_k^2, \quad w_k \notin L_k^1 \cdot L_k^2$$

となる 無限列 $w_0, w_1, \cdots$ と $L_1^1 \cdot L_1^2, L_2^1 \cdot L_2^2, \cdots$ (ただし, $L_i^1 \in \mathcal{L}_1, L_i^2 \in \mathcal{L}_2$) が存在すると仮定する．次のような方法で k_n を帰納的に構成する:

step 0:

w_0 の 2 分割の仕方は有限であり, 任意の $k \geq 1$ に対して, $w_0 \in L_k^1 \cdot L_k^2$ だから, $w_0 = w_0^1 w_0^2$ で $\#\{k \in N \mid w_0^1 \in L_k^1, w_0^2 \in L_k^2\} = \infty$ となる w_0^1, w_0^2 が存在する.

$$N_0 = \{k \in N \mid w_0^1 \in L_k^1, w_0^2 \in L_k^2\}$$
$$k_1 = \min\{k \mid k \in N_0\}$$

とし, step 1 へ行く.

step n $(n > 0)$:

w_{k_n} の 2 分割の仕方は有限であり, 任意の $k \in N_{n-1}$ (ただし, $k \neq k_n$) に対して, $w_{k_n} \in L_k^1 \cdot L_k^2$ だから, $w_{k_n} = w_{k_n}^1 w_{k_n}^2$ で $\#\{k \in N_{n-1} \mid w_{k_n}^1 \in L_k^1, w_{k_n}^2 \in L_k^2\} = \infty$ を満たす $w_{k_n}^1$, $w_{k_n}^2$ が存在する.

$$N_n = \{k \in N_{n-1} \mid w_{k_n}^1 \in L_k^1, w_{k_n}^2 \in L_k^2\}$$
$$k_{n+1} = \min\{k \mid k \in N_n\}$$

とし, step $n+1$ へ行く.

任意の $n \geq 1$ に対して, N_n の定め方から, 明らかに $N_{n-1} \supseteq N_n$ である. したがって, $k_n \leq k_{n+1}$ である. 一方, $k_n \in N_{n-1}$, $k_{n+1} \in N_n$ および, $w_{k_n} \notin L_{k_n}^1 \cdot L_{k_n}^2$ より, $k_n \notin N_n$ である. ゆえに, $k_n < k_{n+1}$ が成り立つ.

この方法で得られる無限列 $w_{k_0}^1 w_{k_0}^2, w_{k_1}^1 w_{k_1}^2, \cdots$; $L_{k_1}^1 \cdot L_{k_1}^2, L_{k_2}^1 \cdot L_{k_2}^2, \cdots$ (ただし, $k_0 = 0$) は, 任意の $n \in N$ に対して,

$$\{w_{k_0}^1, \cdots, w_{k_{n-1}}^1\} \in L_{k_n}^1, \quad \{w_{k_0}^2, \cdots, w_{k_{n-1}}^2\} \in L_{k_n}^2, \quad w_{k_n}^1 w_{k_n}^2 \notin L_{k_n}^1 \cdot L_{k_n}^2$$

を満たしている. $w_{k_n}^1 \notin L_{k_n}^1$ か $w_{k_n}^2 \notin L_{k_n}^2$ の少なくとも一方は成り立つ. すなわち, $w_{k_n}^1 \notin L_{k_n}^1$ を満たす n が無限個存在するか $w_{k_n}^2 \notin L_{k_n}^2$ を満たす n が無限個存在する. 一般性を失うことなく $w_{k_n}^1 \notin L_{k_n}^1$ を満たす n が無限個存在するであるとしてよい. そのような添字をあらためて $k_0, k_1, \cdots$ とすると, 無限列 $w_{k_0}^1, w_{k_1}^1, \cdots$; $L_{k_1}^1, L_{k_2}^1, \cdots$ が存在し, これは $\mathcal{L}_1$ が有限の弾力性をもつことに矛盾する. ■

定義 2.7　　言語族 $\mathcal{L}$ に対して, 次の演算を導入する:

$$\mathcal{L}^{\tilde{n}} = \{L^n \mid L \in \mathcal{L}\}\ (n \geq 1), \quad \mathcal{L}^{\tilde{*}} = \{L^* \mid L \in \mathcal{L}\}, \quad \mathcal{L}^{\tilde{+}} = \{L^+ \mid L \in \mathcal{L}\}$$

(注 $\mathcal{L} \tilde{\cdot} \mathcal{L}$ は $\mathcal{L}^{\tilde{2}}$ とは異なる).

定理 2.5　　言語族 $\mathcal{L}$ が有限の弾力性をもつならば, 言語族 $\mathcal{L}^{\tilde{n}}$, $\mathcal{L}^{\tilde{+}}$, $\mathcal{L}^{\tilde{*}}$ は有限の弾力性をもつ.

proof. これらの証明は定理 2.4 の証明と同様の方法で示すことができる. ■

言語族 $\mathcal{L}$ に対して, $\mathcal{L}^C = \{L^C \mid L \in \mathcal{L}\}$ とする. 次の例は $\boldsymbol{L}$ が言語族の補集合演算に関しては閉じていないことを示す:

例 2.1 Σ^* の帰納的枚挙を $w_0, w_1, \cdots$ とし, $L_k = \Sigma^* - \{w_0, \cdots, w_{k-1}\}(k \in N)$ とする. $\mathcal{L} = L_1, L_2, \cdots$ は有限の厚さであるが, $\mathcal{L}^C$ では任意の $k \in N$ に対して, $\{w_0, \cdots, w_{k-1}\} \subseteq L_k^C$, $w_k \notin L_k^C$ を満たす無限列 $w_0, w_1, \cdots$; $L_1^C, L_2^C, \cdots$ が存在する. したがって, $\mathcal{L}^C$ は有限の弾力性をもたない.

定理 2.6 有限の弾力性をもつ言語族に対して, 演算 $\tilde{\cup}$, $\tilde{\cap}$, $\tilde{\cdot}$, $\cup$, $\tilde{^n}$, $\tilde{^*}$, $\tilde{^+}$ を有限回適用して得られる言語族は正データから帰納推論可能である.

3 EFS の帰納推論

3.1 EFS

Σ を定数記号の有限集合, X を変数記号の可算集合, Π を述語記号の有限集合とする (ただし, Σ, X, Π は互いに素な集合とする). 述語記号は引数の個数を表す非負整数を伴っている. $(\Sigma \cup X)^+$ の要素をパターンという. $\pi_1, \cdots, \pi_n$ をパターン, p を n 引数の述語記号とする. $p(\pi_1, \cdots, \pi_n)$ の形の式をアトムという. パターン π の長さを $|\pi|$ で表し, アトム $p(\pi_1, \cdots, \pi_n)$ に対しては, $|p(\pi_1, \cdots, \pi_n)| = |\pi_1| + \cdots + |\pi_n|$ とする. $A, B_1, \cdots, B_n (n \geq 0)$ をアトムとする. $A \leftarrow B_1, \cdots, B_n$ の形の式を確定節という. そして, 確定節の有限集合を EFS(elementary formal system) という.

節 C は EFS Γ から証明可能である ($\Gamma \vdash C$ で表す) とは, C が Γ の公理から代入と modus ponens を有限回適用することにより得られることをいう (ただし, 代入とは変数記号をパターンに置き換えるもののことをいう).

基礎アトム全体の集合をエルブラン基底といい, HB で表す. エルブラン基底の部分集合 $S \subseteq HB$ をエルブラン解釈という. EFS Γ のすべての節を真にするエルブラン解釈 S をエルブランモデルといい, その中で最小のモデルを最小モデルという ($M(\Gamma)$ で表す). EFS Γ と n 引数の述語記号 p に対して,

$$L(\Gamma, p) = \{(w_1, \cdots, w_n) \in (\Sigma^+)^n \mid \Gamma \vdash p(w_1, \cdots, w_n)\}$$

と定義する．p が 1 引数のとき，$L(\Gamma, p)$ は Σ 上の言語となる．ある言語 L に対し，$L = L(\Gamma, p)$ となる EFS Γ および 1 引数 $p \in \Pi$ が存在するなら L は EFS 言語であるという．また，一般性を失うことなく各言語を定義する 1 引数の述語記号を $p \in \Pi$ と固定してもよい．したがって，$L(\Gamma, p)$ を単に $L(\Gamma)$ とかく．

定義 3.1　節 $A \leftarrow B_1, \cdots, B_n$ が任意の代入 θ に対して，$|A\theta| \geq |B_1\theta| + \cdots + |B_n\theta|$ となるとき，この節は長さ限定であるという．EFS Γ が長さ限定であるとは，任意の節 $C \in \Gamma$ が長さ限定であることをいう．

言語 $L \subseteq \Sigma^+$ が長さ限定 EFS で定義可能であることと L が文脈依存言語であることは同値であることが知られている [2]．また，長さ限定 EFS ではその最小モデルが帰納的であることも示されている [2]．

Shinohara [5] は EFS の帰納推論で重要な役割を果たす次のような既約という概念を導入した．

定義 3.2　EFS Γ が解釈 $S \subseteq HB$ に関して既約であるとは，$S \subseteq M(\Gamma)$ かつ，任意の $\Gamma' \subsetneq \Gamma$ に対して，$S \not\subseteq M(\Gamma')$ が成り立つことをいう．

空でない有限なエルブラン解釈に関して既約な長さ限定 EFS は有限個しか存在しない [5].

3.2　最大長限定 EFS

本節では長さ限定 EFS より構文制約の少ない最大長限定 EFS を定義し，§2.2 で扱った言語族の演算に対応する EFS の演算を最大長限定 EFS の枠組みで表現する．

定義 3.3　節 $A \leftarrow B_1, \cdots B_n$ が任意の代入 θ と任意の $i = 1, 2, \cdots, n$ に対して，$|A\theta| \geq |B_i\theta|$ となるとき，この節は最大長限定であるという．EFS Γ が最大長限定であるとは，任意の節 $C \in \Gamma$ が最大長限定であることをいう．

注) 上記の概念は，Yamamoto [8] が weakly reducing EFS として既に定義されているが本論文で扱う既約な EFS という用語と重複するため，上記の名前を採用した．

定理 3.1　空でない有限なエルブラン解釈に関して，既約な最大長限定 EFS は有限個しか存在しない．

proof. 証明は長さ限定 EFS に関する [5] の定理と同様である. ■

定理 3.2 EFS Γ が最大長限定であるならば, $M(\Gamma)$ は帰納的である.

長さ限定 EFS は最大長限定 EFS であることから, 最大長限定 EFS 言語からなる言語族は文脈依存言語をすべて含む. 文脈依存言語族が, 言語の和, 共通部分, 連接等の演算に関して閉じていことはよく知られている. この節で導入した最大長限定 EFS 言語もまた, これらの演算に関して閉じており, さらにそれらの演算で得られる言語をもとの最大長限定 EFS から簡単に生成できる特徴をもっている. 特に, 共通部分を考えるとき下記の例のように最大長限定 EFS は非常に便利である.

例 3.1 $\Sigma = \{a, b\}$, EFS $\Gamma_1 = \{p(axb)\}$, $\Gamma_2 = \{p(xaby)\}$ とする. $L(\Gamma_1, p) \cap L(\Gamma_2, p)$ は長さ限定で表すことができるが, それはもとの EFS Γ_1, Γ_2 から簡単に構成できない. 次の Γ は $L(\Gamma_1, p) \cap L(\Gamma_2, p) = \{auabvb \mid u, v \in \Sigma^*\}$ から構成した長さ限定 EFS である:

$$\Gamma = \{q(ab)\ ;\ q(xy) \leftarrow q(x)\ ;\ q(xy) \leftarrow q(y)\ ;\ p(axb) \leftarrow q(x)\}$$

しかし, 最大長限定 EFS で考えると, 次のように変数記号の付け替えと 1 つの節を加えることによりもとの EFS から簡単に構成できる:

$$\Gamma = \{p_1(axb)\ ;\ p_2(xaby)\ ;\ p(x) \leftarrow p_1(x), p_2(x)\}$$

次に与えられた最大長限定 EFS 言語の和や共通部分等の言語に対応する EFS の演算を定義する:

定義 3.4 最大長限定 EFS Γ_1, Γ_2 に対して, 次の条件を満たすように述語記号を付け替えた EFS を Γ'_1, Γ'_2 とする: $L(\Gamma_1, p) = L(\Gamma'_1, p_1)$, $L(\Gamma_2, p) = L(\Gamma'_2, p_2)$, Γ'_1, Γ'_2 は述語記号 p を含まずかつ Γ'_1, Γ'_2 の述語記号は互いに素

最大長限定 EFS に対する 2 項演算を次のように定義する:

1. $\Gamma_1 \tilde{\cup} \Gamma_2 = \Gamma'_1 \cup \Gamma'_2 \cup \{p(x) \leftarrow p_1(x)\ ;\ p(x) \leftarrow p_2(x)\}$
2. $\Gamma_1 \tilde{\cap} \Gamma_2 = \Gamma'_1 \cup \Gamma'_2 \cup \{p(x) \leftarrow p_1(x), p_2(x)\}$
3. $\Gamma_1 \tilde{\cdot} \Gamma_2 = \Gamma'_1 \cup \Gamma'_2 \cup \{p(xy) \leftarrow p_1(x), p_2(y)\}$

最大長限定 EFS Γ に対して, 次の条件を満たすように述語記号を付け替えた EFS を Γ とする: $L(\Gamma, p) = L(\Gamma', p')$ かつ Γ' は述語記号 p を含まない

最大長限定 EFS に対する単項演算を次のように定義する:

$$
\begin{array}{llcl}
4. & \Gamma^{\tilde{n}} & = & \Gamma' \cup \{p(x_1 x_2 \cdots x_n) \leftarrow p'(x_1), p'(x_2), \cdots p'(x_n)\} \\
5. & \Gamma^{\tilde{+}} & = & \Gamma' \cup \{p(x) \leftarrow p'(x)\ ;\ p(xy) \leftarrow p(x), p(y)\}
\end{array}
$$

これらの定義から明らかに次のことが成り立つ:

$$
\begin{array}{rclcrcl}
L(\Gamma_1 \,\tilde{\cup}\, \Gamma_2) & = & L(\Gamma_1) \cup L(\Gamma_2) & \qquad & L(\Gamma^{\tilde{n}}) & = & L(\Gamma)^n \\
L(\Gamma_1 \,\tilde{\cap}\, \Gamma_2) & = & L(\Gamma_1) \cap L(\Gamma_2) & & L(\Gamma^{\tilde{+}}) & = & L(\Gamma)^+ \\
L(\Gamma_1 \,\tilde{\cdot}\, \Gamma_2) & = & L(\Gamma_1) \cdot L(\Gamma_2) & & & &
\end{array}
$$

また, 最大長限定 EFS の族 $\mathcal{G}$, $\mathcal{G}_1$, $\mathcal{G}$ に対する演算を次のように定義する:

$$
\begin{array}{rclcrcl}
\mathcal{G}_1 \,\tilde{\cup}\, \mathcal{G}_2 & = & \{\Gamma_1 \,\tilde{\cup}\, \Gamma_2 \mid \Gamma_1 \in \mathcal{G}_1, \Gamma_2 \in \mathcal{G}_2\} & \qquad & \mathcal{G}^{\tilde{n}} & = & \{\Gamma^{\tilde{n}} \mid \Gamma \in \mathcal{G}\} \\
\mathcal{G}_1 \,\tilde{\cap}\, \mathcal{G}_2 & = & \{\Gamma_1 \,\tilde{\cap}\, \Gamma_2 \mid \Gamma_1 \in \mathcal{G}_1, \Gamma_2 \in \mathcal{G}_2\} & & \mathcal{G}^{\tilde{+}} & = & \{\Gamma^{\tilde{+}} \mid \Gamma \in \mathcal{G}\} \\
\mathcal{G}_1 \,\tilde{\cdot}\, \mathcal{G}_2 & = & \{\Gamma_1 \,\tilde{\cdot}\, \Gamma_2 \mid \Gamma_1 \in \mathcal{G}_1, \Gamma_2 \in \mathcal{G}_2\} & & & &
\end{array}
$$

EFS の族 $\mathcal{G}$ が定義する言語族を $L(\mathcal{G}) = \{L(\Gamma) \mid \Gamma \in \mathcal{G}\}$ とかく.

EFS の族に関する演算はそれらが定義する言語族に関する (§2.2 で述べた) 演算に対応する. すなわち, 次のことが成り立つ:

$$
\begin{array}{rclcrcl}
L(\mathcal{G}_1 \,\tilde{\cup}\, \mathcal{G}_2) & = & L(\mathcal{G}_1) \,\tilde{\cup}\, L(\mathcal{G}_2) & \qquad & L(\mathcal{G}^{\tilde{n}}) & = & L(\mathcal{G})^{\tilde{n}} \\
L(\mathcal{G}_1 \,\tilde{\cap}\, \mathcal{G}_2) & = & L(\mathcal{G}_1) \,\tilde{\cap}\, L(\mathcal{G}_2) & & L(\mathcal{G}^{\tilde{+}}) & = & L(\mathcal{G})^{\tilde{+}} \\
L(\mathcal{G}_1 \,\tilde{\cdot}\, \mathcal{G}_2) & = & L(\mathcal{G}_1) \,\tilde{\cdot}\, L(\mathcal{G}_2) & & & &
\end{array}
$$

また, 明らかに $L(\mathcal{G}_1 \cup \mathcal{G}_2) = L(\mathcal{G}_1) \cup L(\mathcal{G}_2)$ である.

§2.2 の定理 2.6 より, 最大長限定 EFS の族の演算に関して次の結果が成り立つ:

定理 3.3 $\mathcal{G}, \mathcal{G}_1, \mathcal{G}_2$ を帰納的枚挙可能な最大長限定 EFS の族とする. 言語族 $L(\mathcal{G})$, $L(\mathcal{G}_1)$, $L(\mathcal{G}_2)$ が有限の弾力性をもつならば, 言語族 $L(\mathcal{G}_1 \,\tilde{\cup}\, \mathcal{G}_2)$, $L(\mathcal{G}_1 \,\tilde{\cap}\, \mathcal{G}_2)$, $L(\mathcal{G}_1 \,\tilde{\cdot}\, \mathcal{G}_2)$, $L(\mathcal{G}_1 \cup \mathcal{G}_2)$, $L(\mathcal{G}^{\tilde{n}})$ および $L(\mathcal{G}^{\tilde{+}})$ はすべて正データから帰納推論可能である.

3.3 有限の弾力性をもつ最大長限定 EFS 言語族

この節では最大長限定 EFS 言語族が有限の弾力性をもつための必要十分条件について論じる. この節で扱う族 $\mathcal{G}$ は, 任意の $\Gamma \in \mathcal{G}$ と任意の $\Gamma' \subseteq \Gamma$ に対して, $\Gamma' \in \mathcal{G}$ を満たすものとする. 集合 $S \subseteq \Sigma^+$ と述語記号 p に対して, $p(S) = \{p(w) \mid w \in S\}$ とする.

定理 3.4　最大長限定 EFS の族を $\mathcal{G}$ とする．言語族 $L(\mathcal{G})$ が有限の弾力性をもつ必要十分条件は，次の (1), (2) を満たす空でない有限集合の無限列 $(T_i)_{i\in N}$ および EFS の無限列 $(\Gamma_i)_{i\in N}$ が存在しないことである (ただし，$T_i \subseteq \Sigma^+$, $\Gamma_i \in \mathcal{G}$):

$$(1)\ \Gamma_1 \subsetneqq \Gamma_2 \subsetneqq \cdots, \quad (2)\ \text{任意の } i \in N \text{ に対して}, \Gamma_i \text{ が } p(T_i) \text{ に関して既約}$$

proof. 必要性) (1), (2) を満たす無限列 $(T_i)_{i\in N}, (\Gamma_i)_{i\in N}$ が存在すると仮定する．まず，$i<j$ ならば，$L(\Gamma_i) \subsetneqq L(\Gamma_j)$ を示す．仮定から $\Gamma_i \subsetneqq \Gamma_j$ であり，Γ_j が T_j に関して既約であることから，$T_j \subseteq L(\Gamma_j)$ かつ $T_j \not\subseteq L(\Gamma_i)$ である．したがって，$L(\Gamma_i) \subsetneqq L(\Gamma_j)$．このことから，この無限列 $(\Gamma_i)_{i\in N}$ は $L(\Gamma_1) \subsetneqq L(\Gamma_2) \subsetneqq L(\Gamma_3) \subsetneqq \cdots$ を満たしている．

$$w_0 \in L(\Gamma_1), \qquad w_i \in L(\Gamma_{i+1}) - L(\Gamma_i) \quad (i \in N)$$

とすると，$w_0, w_1, \cdots; L(\Gamma_1), L(\Gamma_2), \cdots$ は任意の $k \in N$ に対して，

$$\{w_0, w_1, \cdots w_{k-1}\} \subseteq L(\Gamma_k), \quad w_k \notin L(\Gamma_k)$$

を満たし $L(\mathcal{G})$ が有限の弾力性をもつことに矛盾する．

十分性) 任意の $k \in N$ に対して，$\{w_0, w_1, \cdots, w_{k-1}\} \subseteq L(\Gamma_k), w_k \notin L(\Gamma_k)$ を満たす 2 つの無限列 $w_0, w_1, \cdots; L(\Gamma_1), L(\Gamma_2), \cdots$ が存在すると仮定する．

次のように帰納的に k_n と $\tilde{\mathcal{F}}_n$ を定義する (ただし，$k_0 = 0$):

step n $(n \geq 0)$:

$$\begin{aligned}
\mathcal{F}_n &= \{\Gamma \in \mathcal{G} \mid \Gamma : p(\{w_0, w_1, \cdots, w_{k_n}\}) \text{ に関して既約 }\} \\
\tilde{\mathcal{F}}_n &= \mathcal{F}_n - \{\Gamma \in \mathcal{F}_n \mid w_0, w_1, \cdots \text{ がすべて } L(\Gamma) \text{ に属する }\} \\
k_{n+1} &= \min\{k \mid \text{任意の } \Gamma \in \tilde{\mathcal{F}}_n \text{ に対して}, \{w_0, w_1, \cdots, w_k\} \not\subseteq L(\Gamma)\}
\end{aligned}$$

とし，step $n+1$ へ行く．

上記で定義される $k_n, \tilde{\mathcal{F}}_n$ に関して次のことが成り立つ:

(i) 任意の $n \geq 0$ に対して，$\tilde{\mathcal{F}}_n \neq \phi$, $k_n < k_{n+1}$ であること

任意の $k \in N$ に対して，$w_0 \in L(\Gamma_k)$ より，$\Gamma_k \in \mathcal{F}_0$ であるか，または $\Gamma_k \notin \mathcal{F}_0$ ならば $\Gamma'_k \subsetneqq \Gamma_k$ となる $\Gamma'_k \in \mathcal{F}_0$ が存在する．また，$w_k \notin L(\Gamma_k)$ より，後者についても $w_k \notin L(\Gamma'_k)$ であるから，$\tilde{\mathcal{F}}_0 \neq \phi$ が成り立つ．定理 3.1 より，$p(\{w_0\})$ に既約な最大長限定 EFS の個数

は有限であり, また $\tilde{\mathcal{F}}_0$ の定義から, k_1 は定まり, $k_0(=0)<k_1$ である. 任意の $n\in N$ に対して, $\tilde{\mathcal{F}}_n\neq\phi$, および $k_n<k_{n+1}$ であることも同様に示される.

(ii) 任意の $n\geq 0$, 任意の $\Gamma\in\tilde{\mathcal{F}}_{n+1}$ に対して, $\Gamma'\subsetneq\Gamma$ となる $\Gamma'\in\tilde{\mathcal{F}}_n$ が存在すること

k_{n+1} の定め方から, 任意の $\Gamma'\in\tilde{\mathcal{F}}_n$ に対して, $\{w_0,w_1,\cdots,w_{k_{n+1}}\}\not\subseteq L(\Gamma')$ である. したがって, $\Gamma'\notin\tilde{\mathcal{F}}_{n+1}$ となるので $\tilde{\mathcal{F}}_n\cap\tilde{\mathcal{F}}_{n+1}=\phi$. 一方, 任意の $\Gamma\in\tilde{\mathcal{F}}_{n+1}$ に対して, $\{w_0,w_1,\cdots,w_{k_n}\}\subseteq L(\Gamma)$ であるから, $\Gamma'\subsetneq\Gamma$ となる $\Gamma'\in\mathcal{F}_n$ が存在する. また, $\{w_0,w_1,\cdots\}\not\subseteq L(\Gamma)$ であるから, $\{w_0,w_1,\cdots\}\not\subseteq L(\Gamma')$. ゆえに, $\Gamma'\in\tilde{\mathcal{F}}_n$.

(iii) $\Gamma_0\subsetneq\Gamma_1\subsetneq\cdots$ となる $\Gamma_i\in\tilde{\mathcal{F}}_i(i\geq 0)$ が存在すること

(ii) より, 任意の $n\in N$ と 任意の $\Gamma_n\in\tilde{\mathcal{F}}_n$ に対して, $\Gamma_0\subsetneq\Gamma_1\subsetneq\cdots\subsetneq\Gamma_n$ を満たす $\Gamma_i\in\tilde{\mathcal{F}}_i$ $(i=0,\cdots,n-1)$ が存在する. $\#\tilde{\mathcal{F}}_0<\infty$ であるから, $\Gamma_0\subsetneq\Gamma_1\subsetneq\cdots$ を満たす $\Gamma_i\in\tilde{\mathcal{F}}_i$ $(i=0,1,\cdots)$ が存在する. また, $T_i=\{w_0,w_1,\cdots,w_{k_i}\}$ とすると, $\tilde{\mathcal{F}}_i$ の定義より, このような Γ_i が $p(\{w_0,\cdots,w_{k_i}\})$ に関して既約である. ■

この定理から, Shinohara [5] による高々 n 個の節をもつ長さ限定 EFS に関する結果が最大長限定 EFS でも成り立つことがわかる.

系 3.5 $\mathcal{G}=\{\Gamma\mid\Gamma:\text{最大長限定 EFS},\#\Gamma\leq n\}$ とする. 言語族 $L(\mathcal{G})$ は有限の弾力性をもつ. したがって, $L(\mathcal{G})$ は正データから帰納推論可能である.

3.4 帰納推論可能性の特徴づけ定理

この節では最大長限定 EFS の言語族に対する正データからの帰納推論可能性を特徴づける定理を与える. 証明等は長さ限定 EFS に関する論文 [3] と同様なので詳細は省く. 本節でも, 最大長限定 EFS の族 $\mathcal{G}$ は部分集合に関して閉じているものとする. すなわち, 任意の $\Gamma\in\mathcal{G}$ と任意の $\Gamma'\subseteq\Gamma$ に対して, $\Gamma'\in\mathcal{G}$ とする.

補題 3.1 最大長限定 EFS の族を $\mathcal{G}$ とする. $L(\mathcal{G})$ は M - 有限の厚さである.

定理 3.6 最大長限定 EFS の族を $\mathcal{G}$ とする. $L(\mathcal{G})$ が正データから帰納推論可能であること, $L(\mathcal{G})$ が ftt をもつこと, $L(\mathcal{G})$ が pftt をもつことは等価である.

定理 3.7 最大長限定 EFS の族を $\mathcal{G}$ とする. $L(\mathcal{G})$ が正データから帰納推論可能であるための必要十分条件は, 任意の $\Gamma\in\mathcal{G}$ に対して, 次の (1) から (5) を満たすΣ^+ の空でない

有限部分集合の無限列 $(T_i)_{i\in N}$ および EFS の無限列 $(\Gamma_i)_{i\in N}$ が存在しないことである (ただし, $\Gamma_i \in \mathcal{G}$): (1) $T_1 \subsetneq T_2 \subsetneq \cdots$, (2) $\bigcup_{i=1}^{\infty} T_i = L(\Gamma)$, (3) $\Gamma_1 \subsetneq \Gamma_2 \subsetneq \cdots$, (4) 任意 $i \in N$ に対して, Γ_i が $p(T_i)$ に関して既約, (5) 任意 $i \in N$ に対して, $L(\Gamma_i) \subsetneq L(\Gamma)$

論文 [3] では, 長さ限定 EFS の最小モデル族の正データからの帰納推論可能性についても論じているが, そこで得られた結果は本稿で導入した最大長限定 EFS においても同様に成立する. それらの結果は紙面の都合上省略する.

参考文献

[1] D. Angluin. *Inductive inference of formal languages from positive data.* Information and Control, **45:**117–135, (1980).

[2] S. Arikawa, T. Shinohara, and A. Yamamoto. *Elemetary formal system as a unifying framework for language learning.* Proc. 2nd Workshop Comput. Learning Theorey, 312–327, (1989).

[3] 佐藤優子, 森山隆史. 正データからの EFS の帰納推論. November (1992). 重点領域研究 (1) 研究会資料.

[4] M. Sato and K. Umayahara. *Inductive inferability for formal languages from positive data.* IEICE Trans. Inf. & Syst., **E75–D(4):**84–92, (1992).

[5] T. Shinohara. *Inductive inference from positive data is powerful.* Proc. 3rd Workshop Comput. Learning Theory, 97–110, (1990).

[6] T. Shinohara. *Inductive inference of monotonic formal systems from positive data.* New Generation Computing, **8:**371–384, (1991).

[7] K. Wright. *Identification of unions of languages drawn from an identifiable class.* Proc. 2nd Workshop on Comput. Learning Theory, 328–333, (1989).

[8] A. Yamamoto. *Elementary formal system as a logic programming language.* Proc. Logic Program. Conf.'89, ICOT, 123–132, (1989).